Nizhny Novgorod State University
Institute of Information Technologies, Mathematics and Mechanics
Department of Computer Software and Supercomputer Technologies

Educational course «Introduction to deep learning using the Intel® neon™ Framework»

Lecture №5
Transfer learning of deep neural networks

Supported by Intel

Kustikova Valentina

Nizhny Novgorod
2018

Content

1 Abstract 3
2 Literature 3
2.1 Books 3
2.2 References 3

1 Abstract

The transfer learning is one of the widely used approaches in solving practical problems using deep learning. The goal of this lecture is to introduce the notion of ***transfer learning*** of deep neural networks [4] and demonstrate, on a practical example, the effectiveness of this approach.

The idea of the approach is to borrow a trained model to solve the target task. If there is some ***original task***, then to solve it you need to collect the training data and to label this data. A similar procedure can be performed for some other ***target task***. If these tasks are related in some way, then it makes sense to use a trained model for solving the original task for the target one. The goal of learning transfer is to accumulate the knowledge necessary to solve the original task, and use them to solve the target one that is close in meaning [5].

According to the general approach, four types of transfer learning are introduced: direct use of the model, use of a deep model structure, use of the model as a fixed method of feature extraction, fine-tuning the parameters of the original model for solving the target task. The first scheme assumes that the target task is a subtask of the original one. In this case, direct use of the trained model is permissible. Obviously, this approach is not always available. The second scheme assumes that the model constructed for solving the original task is trained on the data prepared for the solution of the target task. In this case, the weights of the model are initialized randomly. The idea of the third scheme is to remove from the deep model the classifier (the last fully-connected layers) and to consider the initial part of the network as a method of feature extraction. In this case, instead of the old classifier, you can put a new classifier (for example, another set of fully-connected layers) and train it on the features constructed using the initial part of the network. The fourth approach is similar to the third, the only difference is that after the classifier is replaced, the entire model is trained as a single system. The key question is how to determine when and what type of transfer learning should be used. The lecture reviews the main factors that have a significant impact on the decision being made [6].

The practical application of the transfer of learning is demonstrated on the task of classifying a person's sex from a photograph. The results of experiments are represented. The results prove the prospects of using this approach in practice.

2 Literature

2.1 Books

1. Haykin S. Neural Networks: A Comprehensive Foundation. – Prentice Hall PTR Upper Saddle River, NJ, USA. – 1998.
2. Osovsky S. Neural networks for information processing. – 2002.
3. Goodfellow I., Bengio Y., Courville A. Deep Learning. – MIT Press. – 2016. – [http://www.deeplearningbook.org].
4. Pan S.J., Yang Q. A Survey on Transfer Learning // IEEE Transactions on Knowledge and Data Engineering. – 2010. – Vol. 22, Issue 10. – P.1345-1359.

2.2 References

5. Transfer Learning – Machine Learning's Next Frontier» – [http://ruder.io/transfer-learning].
6. Transfer Learning [http://cs231n.github.io/transfer-learning].